NOAA GEO-IDE / Unified Access Framework Technical meeting

November 16 - 18, 2010

Contents

Discussion of background and goals:

UAF is built on an agile philosophy: where possible, "Don't solve problems. Copy success." In its first year it has been exploiting the opportunities for data integration that have been created by wide-spread use of netCDF-CF for gridded data. UAF has built a distributed "clean THREDDS catalog" of well-formed, (usually) aggregated gridded data. We have hooked this catalog into a range of client applications and data discovery tools and exploited OPeNDAP access to ensure that data discoverability reliably leads to data accessibility in applications of the end user's choosing.

The first important goal of this meeting will be to assess our accomplishments from year one. How can we grow the clean catalog? make it more complete with respect to the contents of the raw catalog? improve the feedback to data suppliers? Further automate the maintenance of the clean catalog? We will look at selected THREDDS sites (e.g. NODC, NOMADS) as test cases for improving our current UAF solutions.

During year two we will be shifting our focus towards observations. Observations are organized into various classes based upon their "sampling geometries" -- time series (T-lines), vertical profiles (Z-lines), cruise tracks (H-lines), (also cruise sections (HZ-planes) and other 2D structures). A single "feature" from any one of these classes has the character of a 1D (degenerate) grid.

At this meeting we will consider the tools and techniques that are available to expand UAF to address these data types. As a first step the techniques that were employed in year one can be applied: each individual time series, profile, etc. regarded as an independent dataset that is visible in the THREDDS catalog. Existing UAF tools will make it possible for users to discover, visualize, subset and download these individual "features". We will outline a plan for this at the meeting.

At the meeting we want to continue this discussion, understanding options and outlining goals, to provide ?higher level? access to observations as well -- approaches that regard an entire collection as a single data object. We will outline milestones for exploring these approaches, capitalizing (as always in UAF) on work that is being made available by others.

Day 1 Nov. 16 (full day)

A review of year 1 approaches and understanding options for year 2.

Morning

Intro and Welcome. Recap of themes (?Don?t solve problems. Copy success.?)

- During Steve's opening presentation there was some discussion about how to advertise/communicate/highlight the successes of the UAF project and approach.
- Ken raised the need to keep in mind how this work can be leveraged by other NOAA projects and programs.
- One mechanism mentioned was to leverage existing trainings or courses (e.g. annual THREDDS, NetCDF, CF workshop held at Unidata).

Review (only) of year 1: accomplishments, remaining challenges, and lessons learned -- short (10 minute) presentations by members involved with the topics below:

- UAF clean THREDDS catalog and LAS
    - ♦ Kevin gave an overview presentation of the THREDDS catalog cleaner and LAS.
- Data access
    - ♦ IDV
        - ◊ Rich gave an overview of IDV and Matlab.
        - ◊ Rich mentioned two possible success stories to highlight: 1) work done by ocean scientist utilizing HF Radar data and 2) Deep Water Horizon oil spill forecasting.
    - ♦ ERDDAP
        - ◊ Roy gave an overview of ERDDAP.
- Data discovery
    - ♦ Metadata,NcISO
        - ◊ Dave gave an overview on data documentation and leveraging information in THREDDS and netCDF-CF files to create ISO metadata.
    - ♦ RAMADDA
        - ◊ Kevin gave a demonstration of the RAMADDA tool for data discovery.
- Home page
    - ♦ Lewis gave a demonstration of the GEO-IDE/UAF website.
    - ♦ Lewis asked for meeting participants to review the list of questions below and to provide feedback to him in order to improve the website:
        - ◊ Website Questions/Comments for Discussion:
            - · Home Page:
                - Is the text on this page compelling?
                - Does the idea of presenting a ?highlighted dataset and provider? make sense?
            - · Access Data Page:
                - What other tools/capabilities should we present on this page?
                - Does it make sense to organize this page into separate sections to provide links to things like: "Applications I can use", "to include "Tasks I can perform", "Writing UAF programs"?
            - · Find Data Page:
                - What other tools/capabilities should we present on this page?
            - · Contribute Data Page:
                - Are the instructions posted on this page sufficient?
            - · Learn About UAF Page:
                - Is the text on this page compelling?
            - · UAF in Action Page:
                - What are other documented examples of how UAF enabled ?good things? (science) to get done?
            - · GEO-IDE Wiki:

- How can we make better use of this wiki?

Afternoon

Extending UAF to handle collections of observations.

- CDM directions
  - John Caron gave a presentation on the future of THREDDS, CDM, and CF.
- USGS presentation
  - Dave gave a presentation on the USGS Center for Integration of Data Analysis and some of their work.
- Data base ideas
  - Bill Howe gave a presentation on potential strategies and approaches for utilizing relational databases for environmental observations.
- ERDDAP
  - Roy gave an overview of some of the other capabilities of ERDDAP.
- GHRSST at NODC
  - Upendra gave a presentation on how NODC is serving the GHRSST datasets via TDS.

Day 2 Nov. 17 (full day)

This is a day of making plans.

Morning

The following list of 1D data collections and gridded datasets was discussed as potential candidates to include in the UAF catalog:

- NDBC time series/TAO/OceanSITES - CF requiring cleanup
  - Individual CF files
    - Need work
  - Moored timeserires
  - Spectral wave data
  - Elevation data
  - ADCP data
- Argo (at NODC)
  - Talk with ken Casey
  - How to fuse with real time GDACS?
- GTSPP (at NODC) ? CF, ERDDAP/ODV
- SAMOS
  - UAF to work with Shawn on CF files
- ?WOCE hydrographic data (not quite CF -- http://cchdo.ucsd.edu/)
  - Steve to contact
- WODB (both pre- and post- interpolation) ? Dapper served?
- tide gauges at U. Hawaii
  - connect with Mark Merrifield
  - possible connection to HIOOS
  - John Mauer to possibly assist
- HF Radar radial data
  - Jim Potemra/IOOS
- SOCAT carbon cruises
  - PMEL to make data available after official publication

- GTOPPS/POST data (tech success story, no public access to data)
- Investigate tagging data from COML (Roy, Mike Feldman)
- Radiosonde (Source TDB) Legacy format
- Climate stations from NCDC/COOP precip data/GHCN/GSOD
- MADIS
- NGDC Ionosonde ? CF
- OSMC XBTs
    - ERDDAP on OSMC
- ICOADS collection
    - Roy to contact Steve Worley
- Stream gauge services
    - NWIS

The following services were discussed to determine their availability for collections of observations:

- ERDDAP
- single netCDF files files using netcdf3 to represent trajectories, timeseries and profiles in already defined standards
- geoJSON
    - add UAF conventions to describe data types
- OpenLayers(WMS services, GEOext)
- Shared vocabularies/units
- THREDDS point obs feature service
    - Task 1: output netCDF CF-1.5
    - Task 2: output ODV TSV
    - Task 3: input ERDDAP sequences connecting to database
- OGC standards for accessing data (WMS, WFS-WXXM, WPS)
- SOS (IOOS/USGS)
    - ODV

Afternoon

Data Discovery

- Ted gave a presentation on data documentation in the UAF project.
- Overview of existing capabilities that may provide useful data discovery services in the short term (preferably data discovery that is linked to data access):
    - RAMADDA
    - ERDDAP (as a generic de-coupled data discovery capability)
    - GI-Cat
    - Native THREDDS (with more viewers and services)
- current metadata issues and targets for improvement
- catalogs and data discovery approaches (including external mechanisms such as GEOSS CSR)

Day 3 Nov. 18 (half day)

Action Items:

- Testimonials of UAF in Action (Promoting UAF concepts)
    - Call Lewis with testimonials
- Documentation (Scott/Julie)
    - 6 month hiatus

        - Schedule UAF telcon reviewing documentation and integrating meeting results
    - Enhancements to clients
        - R
            - Need homepage tutorial
            - ERDDAP
            - Single netCDF CF files
            - ncstream
        - Matlab
            - Single netCDF files
            - ncstream (when available)
            - Possibility to read xml services
            - ERDDAP (mat files, tabledap)
                - IOOS SOS, waterML-1?
            - Able to read geoJSON (Charles?)
            - Discovery via cut and paste
            - Investigate CSW script for matlab
    - Data discovery
        - explore GI-CAT
            - can you query it via a service
        - explore ESRI data portal
    - Observed data access
        - PMEL to install ERDDAP
        - UAF commitment to help data suppliers serve data through ERDDAP understood services
    - Gridded data
        - NOS hydrographic data at NGDC via THREDDS (need IOSP for HDF5 files?) (Dave Fischman at NOS)
        - GODAS (Scott Cross via OceanNOMADS)
        - Reanalysis.org (Scott, Steve, Phil, Jay)
    - Catalog cleaner
    - Collaborations
    - metadata advancement
    - Outreach
        - find interested parties wanting to share data
    - Get someone from NDBC on UAF telcon
        - Rich to visit NDBC with Bob?s code
        - NDBC to install and run ERDDAP
        - Invite someone from NDBC into UAF
        - Aggregating near real time data and NODC archives (Ken Casey)
    - Explore how to best utilize ERDDAP via THREDDS
    - Dean, Steve, Jay find NASA connection points

Kevin

- Install ncIOS at GFDL

Lewis

- Continue refinement of website
- Continue refinement of wiki
- Testimonials documentation of UAF in action

Ted

- Encourage other TDS providers to install ncISO
- Work with TDS to get ncISO into THREDDS (4.3?)
- Work with Roy to provide access to IOS metadata from ERDDAP
- Working with data providers to increase NcML and ISO scores
- All IOS metadata NcISO creates will be in WAFs
- Metadata repositories (national and international)
- Webinars on metadata improvement
- Visit with data providers

Rich

- Working with gi-cat groups to make sure thredds crawler calculates geospatial extents
- Complete wiki entry for using matlab
- Create example CSW request
- Install ERDDAP for ioos modeling testbed
- Visit NDBC to help improve netCDF files and install ERDDAP
- Test radial feature sites
- Test trajectory collection feature type for particle tracking model output
- Working on unstructured grids with Java netCDF

Roland

- Improve LAS client access to new services (ERDDAP, ncstream, etc)
- Webinars on LAS installation and/or F-TDS
- Prototype LAS gridded-insitu comparisons

Dave

- Help Ted!
- Highlight WMS services in UAF working with GEO-ext/Open Layers libraries

Steve

- Raise issue of data discovery attributes with CF group
- Notify people that CF meeting likely to happen in May (Ashville)
- CF discussions on various feature types
- Make sure CJ?s DWH report mentions UAF
- OCO activities(SAMOS, UH, ARGO, OceanSITES)
- Get permission from SOCAT to serve data
- OSMC project plan to include confluence with UAF

Upendra

- Continue working on aggregations
- Work on point data aggregations
- Test ARC-GIS on the geo portal

Dave/Tom

- Work with unidata to develop feature time series standard
- Using ncstream, cdm remote
- Work with Ted on metadata management
- Work with unidata to get geo tiff IOSP into thredds

Jay/Phil

- Set up ERDDAP
- Populate THREDDS server
- Setup ESG node
- Install ncISO on thredds server
- Get nomads into clean catalog

Scott

- Working on building out oceanNOMADS as operational
- Add content to oceanNOMADs
- Set up LAS/F-TDS
- Investigate setting up ERDDAP
- Documentation

Ken

- Unidata relationship for support of UAF needs
- Identify someone at NDBC to engage with UAF
- Outreach for NexGen 4-d web cube

Jeff

- Maintain connection between UAF and weather data cube activity

Matt

- Help Lewis with testimonials
- Outreach to GIS community
- Explore GI-GO/GI-CAT
- Explore connection of spatial database to thredds
- Work on shared vocabularies

Charles

- Connect with NDBC thredds contact to fix up server
- Keep current thredds setup uaf compliant
- Get RTOFS served as grib2
- Update thredds server to 4.2
- Aggregate higher resolution NCOM American seas data
- Install NcISO on edac thredds
- Use thredds point obs when available
- Help Scott with documentation

Roy

- ICOADS - both gridded and observational data
- WODB - work with Upendra
- ERDDAP
    - ♦ Do tutorials on using ERDDAP itself
    - ♦ Develop scripts for Matlab and R
    - ♦ Provide support on installation to other groups

- Look at adding WaterML services to ERDDAP
- See if we can get it so that ERDDAP can hook up to OSMC database, return files in desired formats

Meeting Attendees:

Name	Organization	E-mail
Matt Austin	NOAA/NESDIS	Matthew.Austin@noaa.gov
Dave Blodgett	USGS	dblodgett@usgs.gov
James Calkins	NOAA/NWS	James.Calkins@noaa.gov
Charles Carleton	NOAA/NESDIS	Charles.Carleton@noaa.gov
John Caron	Unidata	caron@unidata.ucar.edu
Phil Cogbill	NOAA/NESDIS	Phil.Cogbill@noaa.gov
Scott Cross	NOAA/NESDIS	Scott.Cross@noaa.gov
Upendra Dadi	NOAA/NESDIS	Upendra.Dadi@noaa.gov
Ted Habermann	NOAA/NESDIS	Ted.Habermann@noaa.gov
Steve Hankin	NOAA/OAR	Steven.C.Hankin@noaa.gov
Jay Hnilo	NOAA/NESDIS	Jay.Hnilo@noaa.gov
Bill Howe	University of Washington	billhowe@cs.washington.edu
Heather Koyuk	NOAA/OAR	Heather.Koyuk@noaa.gov
Tom Kunicki	USGS	tkunicki@usgs.gov
Lewis McCulloch	NOAA/NESDIS	Lewis.Mcculloch@noaa.gov
Ken McDonald	NOAA/NESDIS	Kenneth.Mcdonald@noaa.gov
Roy Mendelssohn	NOAA/NMFS	Roy.Mendelssohn@noaa.gov
Dave Neufeld	NOAA/NESDIS	David.Neufeld@noaa.gov
Kevin O'Brien	NOAA/OAR	Kevin.M.O'Brien@noaa.gov
Roland Schweitzer	NOAA/OAR	Roland.Schweitzer@noaa.gov
Rich Signell	USGS	rsignell@usgs.gov
Jeff Smith	NOAA/NWS	Jeff.S.Smith@noaa.gov